Exercises Econometrics: Set 1
Due December 5th, 2013.

General guidelines:

Make sure to be able to carefully explain your answers and interpret the results. The objective of this class is not only to understand and perform data analysis using different econometric techniques, but also to help you interpret the results of empirical studies. It is the interpretation that is often the most interesting part of empirical work, in particular if we want to derive policy implications from our findings.

Computer exercises:

1) Use the data in SLEEP75.RAW from Biddle and Hamermesh (1990) to study whether there is a tradeoff between the time spent sleeping per week and the time spent in paid work. We could use either variable as the dependent variable.
For concreteness, estimate the model $sleep = \beta_0 + \beta_1 totwork + u$, where *sleep* is minutes spent sleeping at night per week and *totwrk* is total minutes worked during the week.

(i) Report your results in equation form along with the number of observations and $R2$. What does the intercept in this equation mean?
(ii) If *totwrk* increases by 2 hours, by how much is *sleep* estimated to fall? Do you find this to be a large effect?
(iii) Let *totwrk_hr* be the total hours worked during the week. Without using the computer, what will be the coefficient of this variable in a regression of sleep on *totwrk_hr*?

2) A more realistic version of this model is: $sleep = \beta_0 + \beta_1 totwork + \beta_2 educ + \beta_3 age + u$,

where *sleep* and *totwrk* (total work) are measured in minutes per week and *educ* and *age* are measured in years.

(i) If adults trade off sleep for work, what is the sign of β_1?
(ii) What signs do you think β_2 and β_3 will have?
(iii) Using the data in SLEEP75.RAW, estimate the equation above. Write down the equation with the found coefficients, the R^2, and the number of observations.
(iv) If someone works five more hours per week, by how many minutes is *sleep* predicted to fall? Is this a large tradeoff? Compare these results with the results you had in the first exercise.
(v) Discuss the sign and magnitude of the estimated coefficient on *educ*.
(vi) Would you say *totwrk*, *educ*, and *age* explain much of the variation in *sleep*? What other factors might affect the time spent sleeping? Are these likely to be correlated with *totwrk*?

Exercise without computer:

3) The median starting salary for new law school graduates is determined by $\log(salary) = \beta_0 + \beta_1 LSAT + \beta_2 GPA + \beta_3 \log(libvol) + \beta_4 \log(\cos t) + \beta_5 rank + u$, where *LSAT* is median LSAT score for the graduating class, *GPA* is the median college GPA for the class, *libvol* is the number of volumes in the law school library, *cost* is the annual cost of attending law school, and *rank* is a law school ranking (with *rank*=1 being the best).

(i) Explain why we expect $\beta_5 \le 0$.

(ii) What signs to you expect for the other slope parameters? Justify your answers.

(iii) Using the data in LAWSCH85.RAW, the estimated equation is:

$$\log(salary) = 8.34 + 0.0047 LSAT + 0.248 GPA + 0.095 \log(libvol) + 0.038 \log(\cos t) - .0033 rank + u,$$

$$n = 136,\ R^2 = 0.842.$$

What is the predicted ceteris paribus difference in salary for schools with a median GPA different by one point? (Report your answer as a percent.)

(iv) Interpret the coefficient on the variable log(*libvol*).

(v) Would you say it is better to attend a higher ranked law school? How much is a difference in ranking of 20 worth in terms of predicted starting salary?